AD 741803

ARL 72-0033
FEBRUARY 1972

Aerospace Research Laboratories

NONNULL DISTRIBUTION OF HOTELLING'S GENERALIZED T_0^2 STATISTIC

A. K. CHATTOPADHYAY
APPLIED MATHEMATICS RESEARCH LABORATORY

PROJECT NO. 7071

Approved for public release; distribution unlimited.

AIR FORCE SYSTEMS COMMAND
United States Air Force

21

NOTICES

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely related Government procurement operation, the United States Government thereby incurs no responsibility nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

Agencies of the Department of Defense, qualified contractors and other government agencies may obtain copies from the

Defense Documentation Center
Cameron Station
Alexandria, Virginia 22314

This document has been released to the

CLEARINGHOUSE
U. S. Department of Commerce
Springfield, Virginia 22151

for sale to the public.

Copies of ARL Technical Documentary Reports should not be returned to Aerospace Research Laboratories unless return is required by security considerations, contractual obligations or notices on a specified document.

72/500

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Aerospace Research Laboratories Applied Mathematics Research Laboratory Wright-Patterson AFB, Ohio 45433	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

Nonnull Distribution of Hotelling's Generalized T_0^2 Statistic

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Scientific Interim

5. AUTHOR(S) (First name, middle initial, last name)

A. K. Chattopadhyay

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
February 1972	19	7

8a. CONTRACT OR GRANT NO. In-house Research b. PROJECT NO. 7071-00-12 c. DoD Element 61102F d. DoD Subelement 681304	9a. ORIGINATOR'S REPORT NUMBER(S) 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) ARL 72-0033

10. DISTRIBUTION STATEMENT

Approved for public release; distribution unlimited.

11. SUPPLEMENTARY NOTES TECH OTHER	12. SPONSORING MILITARY ACTIVITY Aerospace Research Laboratories (LB) Wright-Patterson AFB, Ohio 45433

13. ABSTRACT

In ~~this~~ paper the author derived asymptotic expressions for percentile and c.d.f. of Hotelling's Generalized T_0^2 statistic under the nonnull assumption of mean matrix and variance covariance matrix satisfy (3) and (4) given in the text.

These expressions can be used to study the robustness of the test with respect to power function and stabilization of critical region.

DD FORM 1 NOV 65 1473

UNCLASSIFIED
Security Classification

UNCLASSIFIED

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
nonnull distribution						
percentile, c.d.f.						
Hotelling's T_0^2 statistic						

☆U.S.Government Printing Office: 1972 — 759-085/530

UNCLASSIFIED

Security Classification

ARL 72-0033

NONNULL DISTRIBUTION OF HOTELLING'S GENERALIZED T_0^2 STATISTIC

A. K. CHATTOPADHYAY

APPLIED MATHEMATICS RESEARCH LABORATORY

FEBRUARY 1972

PROJECT 7071

AEROSPACE RESEARCH LABORATORIES
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
WRIGHT-PATTERSON AIR FORCE BASE, OHIO

FOREWORD

This report was prepared for Applied Mathematics Research Laboratory, Aerospace Research Laboratories, by A.K. Chattopadhyay under Project 7071, Research in Applied Mathematics. This work was performed at U.S.A.F. Aerospace Research Laboratories by the author while in the capacity of Technology Incorporated Visiting Research Associate under contract F33615-71-C-1463.

In this report the author studies the robustness of Hotelling's Generalized T_0^2 test under the violation of general linear hypothesis both in respect of mean and variance covariance matrices.

The author wishes to thank Dr. P.R. Krishnaiah for some useful discussions, and Mrs. Georgene Graves for typing the manuscript carefully.

ABSTRACT

In this paper the author derived asymptotic expressions for percentile and c.d.f. of Hotelling's Generalized T_0^2 statistic under the nonnull assumption of mean matrix and variance covariance matrix satisfy (3) and (4) given in the text.

These expressions can be used to study the robustness of the test with respect to power function and stabilization of critical region.

TABLE OF CONTENTS

1. Introduction

In an earlier paper [2] the asymptotic formulae for the distribution and percentile of statistic $T = m\,\mathrm{tr}\,S_1 S_2^{-1}$ have been obtained up to terms $\frac{1}{n}$ where mS_1 and nS_2 are independently distributed $W(m,p,\Sigma_1)$ and $W(n,p,\Sigma_2)$ respectively. Similar expansions for the ratio of two independent trace statistics are also obtained. This, in fact, generalizes the work of previous authors [3], [5], [6]. In a recent paper [6] Siotani gave an asymptotic expansion for the nonnull distribution of Hotelling's generalized T_0^2 up to terms $\frac{1}{n^2}$ by using James [4] and Welch [7] idea by expanding the characteristic function by the perturbation technique.

In this article we generalize the earlier results and find asymptotic expansion up to terms of order $\frac{1}{n}$ for c.d.f. and percentile of the trace statistic when mS_1 has $W(m,p,\Sigma,\Omega)$ and $\Sigma_1 \neq \Sigma_2$ but otherwise satisfy (3) and (4). The expression (6) can be used to compute the power of T_0^2 test for the generalized linear hypothesis when departure from hypothesis $M = 0$ and $\Sigma_1 = \Sigma_2$ is present.

It can also be used in cases to test two covariance matrices when one has a noncentral Wishart distribution.

2. Formulation of the Problem

Let $Z = \{z_1, \ldots, z_m\}$ be a $p \times m$ random matrix where z_i's are independently distributed according to p variate normal distribution with mean vector μ_i and variance covariance matrix $\Sigma_1 = B^{-1}$. Let $nS_n = n(s_{ij})$ be a $p \times p$ matrix which is independent of Z and follows a central Wishart distribution $W(n,p,\Lambda^{-1})$ with n degrees of freedom and variance covariance matrix $\Sigma_2 = \Lambda^{-1}$. Hotelling's generalized is given as

$$T_o^2 = \operatorname{tr} S_n^{-1} ZZ'$$

$$= \sum_{i=1}^{m} Z_i' S_n^{-1} Z_i \text{ , when } \Sigma_1 = \Sigma_2$$

Our aim is to find asymptotic expansion for percentile and c.d.f. of T_o^2 when $\Sigma_1 \neq \Sigma_2$ but otherwise satisfying (3).

3. Asymptotic Expansion for Percentile

Let

$$G(\theta) = \Pr\{\mathrm{tr}\, S_n^{-1} ZZ' \le \theta\}$$

Now

$$\Pr\{\mathrm{tr}BZZ' \le \theta\}$$

$$= e^{-\frac{\omega^2}{2}} \sum_{j=0}^{\infty} \frac{\left(\frac{\omega^2}{2}\right)^j}{j!\,2^{\rho+j}\,\Gamma(\rho+j)} \int_0^{\theta} x^{\rho+j-1} e^{-\frac{x}{2}} dx$$

$$= \chi^2_{mp}(\theta, \omega^2)$$

Where

$$\omega^2 = \mathrm{tr}BMM' = \mathrm{tr}\Omega$$

$$M = \{\mu_1, \ldots, \mu_m\} \ne 0, \quad \rho = \frac{mp}{2}$$

and $\chi^2_{mp}(\theta, \omega^2)$ is the c.d.f. of noncentral chi-square variable with mp d.f. and noncentrality parameter ω^2.

Now we can try to find a function $h(S_n)$ of the elements of S_n, n being large enough such that

$$G(\theta) = \Pr\{\mathrm{tr}\, S_n^{-1} ZZ' \le h(S_n)\}$$

Now

$$G(\theta) = E_{S_n} \{Pr\ \{exp\{tr(S_n - A^{-1})\ \partial\}$$

$$Pr\ \{tr\ AZZ' \leq h(A^{-1})\}\}\}$$

$$= H\ Pr\ \{tr\ AZZ' \leq h(A^{-1})\}$$

Where

$$H = exp\ [-tr\ A^{-1}\ \partial]\ \left| I - \frac{2}{n} A^{-1}\ \partial \right|^{-\frac{n}{2}}$$

$$= 1 + \frac{1}{n} \sum_{r,s,t,u} \sigma_{rs}\ \sigma_{tu}\ \partial_{st}\ \partial_{ur} + O(n^{-2}) \qquad (1)$$

$$\partial(p \times p) = (\tfrac{1}{2}\ (1 + \delta_{ij})\ \frac{\partial}{\partial\sigma_{ij}})$$

$$A^{-1} = \begin{pmatrix} \sigma_{11} & & & \sigma_{1p} \\ . & & & . \\ . & & & . \\ . & & & . \\ \sigma_{p1} & . & . & . & \sigma_{pp} \end{pmatrix}$$

Now expanding $h(S_n)$ around θ we get

$$h(S_n) = \theta + h_1(S_n) + h_2(S_n) + \ldots$$

Where $h_i(S_n)$ is $O(n^{-i})$

Thus expanding $h(S_n)$ around $h(A^{-1})$ we get

$$G(\theta) = [1 + \frac{1}{n} \sum \sigma_{rs}\ \sigma_{tu}\ \partial_{st}\ \partial_{ur} + O(n^{-2})]$$

$$[1 + h_1(A^{-1})\, D + O(n^{-2})]\ \Pr\ \{\mathrm{tr}\ AZZ' \leq \theta\ \}$$

Where $D = \frac{\partial}{\partial\theta}$

This being true for all large n we get

$$[h_1\ (A^{-1})\ D + \frac{1}{n} \sum \sigma_{rs}\ \sigma_{tu}\ \partial_{st}\ \partial_{ur}]\ \Pr\ \{\mathrm{tr}\ AZZ' \leq \theta\} = 0$$

Again let

$$J = \Pr\ \{\mathrm{tr}\ (A^{-1} + \varepsilon)^{-1} ZZ' \leq \theta\}$$

Following [2], [3], [6] we get

$$J = |I - x\Delta|^{-\frac{m}{2}}\ \exp(-\frac{\omega^2}{2})$$

$$\exp\ \{\tfrac{1}{2}\ E\ \mathrm{tr}(I - x\Delta)^{-1}\ \Omega\}\ \chi^2_{mp}(\theta,\ 0) \qquad (2)$$

Where

$$\Delta = E - 1, \quad \mathcal{E}^r\ \chi^2_{mp}(\theta,\ \omega^2)$$

$$= \chi^2_{mp + 2r}\ (\theta,\ \omega^2) \text{ and}$$

$$x = B^{-1}(A^{-1} + \varepsilon)^{-1} - I$$

$$= (B^{-1}A - I) - \sum \varepsilon_{rs}\ (B^{-1}A)\ (A^{-1}_{rs}A)$$

$$+ \sum \varepsilon_{rs}\ \varepsilon_{tu}\ (B^{-1}A)\ (A^{-1}_{rs}\ A)\ (A^{-1}_{tu}\ A) - \ldots$$

Where A^{-1}_{rs} is the p x p matrix with (i,j)th element

$$\tfrac{1}{2}(\delta_{ri}\ \delta_{sj} + \delta_{rj}\ \delta_{si})$$

Now let

$$|\text{chi}\ (B^{-1}A - I)| = |\text{chi}\ (F)| < 1,\ i = 1,\ ..,\ p \qquad (3)$$

Where

$$B^{-1}A - I = F$$

Expanding (2) and equating coefficients of $\varepsilon_{rs}\ \varepsilon_{tu}$ with those in Taylor's expansion of J around $\varepsilon = 0$ and denoting

$$\text{tr}\ (A^{-1}{}_{rs}A) = (rs)$$

$$\text{tr}\ (A^{-1}{}_{rs}A)\ (A^{-1}{}_{tu}A) = (rs|tu)$$

$$\text{tr}\ F(A^{-1}{}_{rs}A)\ (A^{-1}{}_{tr}A) = (F|rs|\ tu)$$

$$\text{tr}\ F^2 = (F|F)$$

$$\text{tr}\ F = (F) \qquad \text{etc.}$$

and using the following

$$\sum \sigma_{st}\ \sigma_{ur}\ (rs|tu) = \tfrac{1}{2}p(p + 1)$$

$$\sum \sigma_{rs}\ (rs) = p$$

$$\sum \sigma_{st}\ \sigma_{ur}\ (rs)\ (tu) = p$$

$$\sum \sigma_{st}\ \sigma_{ur}\ (F|rs)\ (tu) = (F)$$

$$\sum \sigma_{st}\ \sigma_{ur}\ (F|rs|tu) = (F)\ (p + 1)\ /2$$

$$\sum \sigma_{st}\, \sigma_{ur}\, (F|rs)\, (F|tu) = (F|F')$$

$$\sum \sigma_{st}\, \sigma_{ur}\, (\Omega|rs|F|tu) = \tfrac{1}{2}[(\Omega)\, (F) + (\Omega F)] \quad \text{etc.}$$

we get after neglecting terms involving $f_{ij} f_{kl}$ when $F = (f_{ij})$ (4)

$$- h_1\, (A^{-1})\, D\, \Pr\, \{\mathrm{tr}\, AZZ' \leq \theta\}$$

$$= \frac{1}{4n} \sum_{j=0}^{4} a_j(m,p)\, g_{mp+2j}(\theta,\, \omega^2)$$

$$+ \frac{1}{4n} \sum_{j=0}^{5} b_j(m,p)\, g_{mp+2j}(\theta,\, \omega^2)$$

$$+ O(n^{-2})$$

Where

$$a_0(m,p) = mp(m-p-1)$$

$$a_1(m,p) = -2m(mp-\omega^2)$$

$$a_2(m,p) = mp(m+p+1) - 2(2m+p+1)\omega^2 + \mathrm{tr}\Omega^2$$

$$a_3(m,p) = 2\{(m+p+1)\omega^2 - \mathrm{tr}\Omega^2\}$$

$$a_4(m,p) = \mathrm{tr}\Omega^2$$

$$b_0(m,p) = -\frac{m^2}{2}\, (F)\, p(m-p-1)$$

$$b_1(m,p) = \frac{m}{2}\, (F)\, \{4m - 3m^2 p - mp^2 - mp - 2(p+m+1)\, (\Omega)\}$$

$$- \frac{m}{2}(\Omega F)\, \{p\, (m-p-1) + 4\}$$

$$b_2(m,p) = -\frac{(F)}{2}\{m^2p^2 + 8m^2 + m^2p + 3m^3p + 4mp + 4m$$

$$- (6m^2 + 8mp + 8m + 4)\ (\Omega)\}$$

$$- \frac{(\Omega F)}{2}\{mp + mp^2 - 3m^2p - 16m - 4p - 8 - (4m + 2p + 2)\ (\Omega)\}$$

$$+ 2m(\Omega F') - 2(\Omega' F\Omega)$$

$$b_3(m,p) = \frac{(F)}{2}\{m^2p^2 + m^2p + 4mp + 4m + m^3p + 4m^2$$

$$- (6m^2 + 10mp + 10m + 8)\ (\Omega)\}$$

$$- \frac{(\Omega F)}{2}\{3m^2p + mp + mp^2 + 2m + 12p + 20$$

$$- (12m + 8p + 8)\ (\Omega)\} + 6(\Omega' F\Omega) + 2(\Omega F\Omega')$$

$$b_4(m,p) = \frac{(F)}{2}\{m^2p + mp^2 + mp + (2m^2 + 4mp + 4m + 4)\ (\Omega)\}$$

$$+ \frac{(\Omega F)}{2}\{8m + 8p + 12 - (10p + 22)\ (\Omega)\}$$

$$+ 2m(\Omega F') - 6(\Omega' F\Omega)$$

$$b_5(m,p) = 2(\Omega F)\ (\Omega)\ (m + P + 1) + 2(\Omega' F\Omega)$$

$$+ 2(\Omega F\Omega')$$

Where as noted earlier we dropped terms involving $f_{ij}f_{kl}$, etc., and $g_{mp}(\theta,\omega^2)$ means the c.d.f. of noncentral chi-square distribution with mp d.f. and noncentrality parameter ω^2.

Thus

$$h(S_n) = \hat{\theta} - [\frac{1}{4n} \sum_{j=0}^{4} a_j(m,p)\, g_{mp+2j}(\hat{\theta},\omega^2)$$
$$+ \frac{1}{4n} \sum_{j=0}^{5} b_j(m,p)\, g_{mp+2j}(\hat{\theta},\omega^2)] \quad [G'(\hat{\theta})]^{-1}$$
$$+ O(n^{-2}) \qquad (5)$$

Where $\hat{\theta}$ is the corresponding percentile of linear function of non-central chi-square variable of the form [1]

$$Y = \sum_{j=1}^{p} \lambda_j \chi_j^2(m)$$

Where λ_j's are the characteristic roots of AB^{-1} and $G(\hat{\theta})$ is the c.d.f. of Y.

Here our form (5) differs slightly from those given in [2], [3], [5], [6] but this form gives a uniform result both for the percentile and for the c.d.f. as given below.

4. Approximation for c.d.f. of Hotelling's Generalized T^2 Under Nonnull Assumptions on Mean Vector and Variance Covariance Matrix

Here we proceed as in [2], [3], [5] and using our earlier calculation we get

$$\Pr\{\operatorname{tr} S_n^{-1} ZZ' \leq \theta\} = H \Pr\{\operatorname{tr} \Lambda ZZ' \leq \theta\}$$

Where H is given by (1).

Thus we get

$$\Pr\{\operatorname{tr} S_n^{-1} ZZ' \leq \theta\} = G(\theta) - \frac{1}{n}[h_1(A^{-1})]\, G'(\theta)$$
$$+ O(n^{-2})$$

on further assumption of (2) and (3) we get

$$\Pr\{\operatorname{tr} S_n^{-1} ZZ' \leq \theta\} = G(\theta) + \frac{1}{4n}\sum_{j=0}^{4} a_j(m,p)$$
$$g_{mp+2j}(\theta,\omega^2) + \frac{1}{4n}\sum_{j=0}^{5} b_j(m,p)\, g_{mp+2j}(\theta,\omega^2)$$
$$+ O(n^{-2}) \tag{6}$$

Where $a_j(m,p)$, $b_k(m,p)$; $j=0, \ldots, 4$, $k=0, \ldots, 5$, $G(\theta)$ and $G_{mp}(\theta,\omega^2)$ are defined earlier.

5. Summary

Summarizing we state the following

Theorem. Let $Z = (z_1, \ldots, z_m)$ be a $p \times m$ random matrix where z_i's are as defined in the formulation and let nS_n be a $p \times p$ matrix which follows a central Wishart distribution $W(n,p,A^{-1})$ independently of Z, then let

$$\text{(i) } B^{-1} A = I + F \text{ and } |\mathrm{ch}_i\,(F)| < 1, \; i=1, \ldots, p$$

(ii) terms involving $f_{ij}f_{kl}$ are negligible where f_{ij} is the (i,j)th element of F.

Then an asymptotic expansion for percentile and c.d.f. of $T = S_n^{-1} ZZ'$ are given by (5) and (6) respectively.

6. Remarks

(a) Putting $F = 0$ in (6) we get the expression due to Siotani [6] up to the indicated order. We note that $b_j(m,p)$ terms vanishes in this case.

(b) Putting $M = 0$ in our model we get the expression given in [2]. This should be immediate if we put $M = 0$ and note that in this case J in (2) reduces to corresponding expression in [2].

REFERENCES

[1] Box, G.E.P. (1954), "Some Theorems on Quadratic Forms Applied in the Study of Analysis of Variance Problems. I. Effect of Inequality of Variance in the One Way Classification", Ann. Math. Statist. 25, 290-302.

[2] Chattopadhyay, A.K. and Pillai, K.C.S. (1971), "Asymptotic Formulae for the Distribution of Some Criteria for Tests of Equality of Covariance Matrices", Journal of Multivariate Analysis. 1, 215-231.

[3] Ito, Koichi (1960), "Asymptotic Formulae for the Distribution of Hotelling's Generalized T_0^2 Statistic II", Ann. Math. Statist. 31, 1148-1153.

[4] James, G.S. (1954), "Tests of Linear Hypothesis in the Univariate and Multivariate Analysis when the Ratios of the Population Variances are Unknown", Biometrika. 41, 19-43.

[5] Siotani, Minoru (1956), "On the Distribution of Hotelling's T^2 Statistic", Ann. Inst. Statist. Math. 8, 1-14.

[6] Siotani, Minoru (1971), "An Asymptotic Expansion of the Non-Null Distribution of Hotelling's T_0^2 Statistic",Ann. Math. Statist. 42, 560-571.

[7] Welch, B.L. (1947), "The Generalization of Students' Problem when Several Different Population Variances are Involved", Biometrika 34, 28-35.